



A Book Review by NATD Member Linda Elman

Jaeger, R.M., and Tucker, C.G. (1998) *Analyzing, Disaggregating, Reporting, and Interpreting Students' Achievement Test Results: A Guide to Practice for Title I and Beyond*, Washington D.C.: Council of Chief State School Officers.

Working with schools both in my district and across the state of Washington, I am excited about the growing interest in using data for exploring student achievement. School personnel are excited about tools that allow them to disaggregate data on various demographic factors, and are eager to know how to work with the data. They frequently ask questions such as “How much difference is worth looking at, what’s the best way to look at these data, or more even commonly, Linda, can you break my data down this way or that way?” So, I was eager to have the opportunity to read and review the Jaeger and Tucker work. The book is short—only 65 pages, and I thought it might make a nice piece to use working with principals in my district. It might give them some of the tools they need to work with their staff in exploring assessment results. In fact, the authors state that the book is written for teachers, principals, and other building, district, and state staff.

At the opening of the document is an index to topics of interest. This index identifies which pages deal with each of the various concepts explored such as “Identifying Questions” or protecting confidentiality by not reporting out groups that are so small that individual identities could be discovered. At the end of the book is a series of appendixes that include directions for computing effect sizes or computing confidence intervals around percentages or averages. The appendixes themselves are very useful.

The central portion of the book is organized around three vignettes that illustrate a school district, a school, and a state exploring their student achievement data in line with and beyond the requirements of the new Title I requirements. Each vignette begins with a set of questions that are addressed through the story. In each one, a committee or team of staff discuss how they are exploring their data, including describing the uses, advantages, and limitations of various graphic and statistical techniques such as using box plots or calculating standard errors to establish confidence intervals as they interpret their results.

Unfortunately, the body of the text is difficult to follow. The text is very dense. It consists of the records of the meetings of the various groups who were analyzing the data—at meeting one this occurred, at meeting two Dr. So and So said . . . , etc. The message of the techniques used, the information gained, and the issues and problems addressed get buried in the prose. With few subheadings and only an occasional bulleted list, the text made it difficult to isolate the issues that are being addressed. And the language, while intended to be conversational, is often very

difficult to understand. For example, one marginal highlight drawn from the text on page 35 reads:

“Knowing that an observed difference is unlikely to be equal to zero in the populations underlying the observed samples offers no more than a small degree of comfort and tells nothing about whether the observed difference is large or small in a substantive sense.”

While this excerpt is a little more intelligible in the context of the paragraph that surrounds it, it is not the kind of text that the principals and teachers I know would be comfortable reading.

Equally frustrating is the placement of figures. Since the graphics are not generally embedded in the sections where they are being discussed, I often found myself leafing back several pages to find the right figure. Further, the figures are frequently placed in clusters that display different disaggregations of the same dependent variable, without explicit explanations near the figures explaining the similarities or differences. I found myself wishing that the authors had used the marginal space (often used for quotations throughout the text) next to the figures to highlight interpretations, or limitations in interpretations for that figure; or comparisons and contrasts among adjacent figures.

Finally, the book only deals with criterion referenced assessment, although I can't find reference to that limitation in the introductory section. This may have kept the book short, but it does not advise those schools, states, and districts working with norm-referenced testing about the difficulties that lie in disaggregating those scores. (So CTBS averages NCE's and ITBS averages *standard scores* which are then converted via a look-up table to percentiles—does that really matter?)

In this day of increased accountability our schools are being encouraged (if not forced) to study their assessment data in ways they have never been asked to do before. We need good tools for test directors that help staff in our buildings look at their data through a variety of lenses. Those districts, too small to have a trained test director, need these kinds of tools even more. The Jaeger and Tucker book, in its current form, is unlikely to serve that need.